Optimized delay of the second COVID-19 vaccine dose reduces ICU admissions

Paulo J. S. Silva[a], Claudia Sagastizábal[a], Luís Gustavo Nonato[b], Claudio José Struchiner[c], and Tiago Pereira[b,1]

[a]Instituto de Matemática, Estatística e Computação Científica, Universidade Estadual de Campinas, 13083-859 São Paulo, Brazil; [b]Instituto de Ciências Matemáticas e Computação, Universidade de São Paulo, 13566-590 São Paulo, Brazil; and [c]Escola de Matemática Aplicada, Fundação Getúlio Vargas, 22250-9 Rio de Janeiro, Brazil

Edited by David L. Donoho, Stanford University, Stanford, CA, and approved July 8, 2021 (received for review March 12, 2021)

Slower than anticipated, COVID-19 vaccine production and distribution have impaired efforts to curtail the current pandemic. The standard administration schedule for most COVID-19 vaccines currently approved is two doses administered 3 to 4 wk apart. To increase the number of individuals with partial protection, some governments are considering delaying the second vaccine dose. However, the delay duration must take into account crucial factors, such as the degree of protection conferred by a single dose, the anticipated vaccine supply pipeline, and the potential emergence of more virulent COVID-19 variants. To help guide decision-making, we propose here an optimization model based on extended susceptible, exposed, infectious, and removed (SEIR) dynamics that determines the optimal delay duration between the first and second COVID-19 vaccine doses. The model assumes lenient social distancing and uses intensive care unit (ICU) admission as a key metric while selecting the optimal duration between doses vs. the standard 4-wk delay. While epistemic uncertainties apply to the interpretation of simulation outputs, we found that the delay is dependent on the vaccine mechanism of action and first-dose efficacy. For infection-blocking vaccines with first-dose efficacy $\geq$50%, the model predicts that the second dose can be delayed by $\geq$8 wk (half of the maximal delay), whereas for symptom-alleviating vaccines, the same delay is recommended only if the first-dose efficacy is $\geq$70%. Our model predicts that a 12-wk second-dose delay of an infection-blocking vaccine with a first-dose efficacy $\geq$70% could reduce ICU admissions by 400 people per million over 200 d.

outbreaks | vaccination | control | strategies

Immunization of a large proportion of the worldwide population against the severe acute respiratory syndrome coronavirus 2 (COVID-19) is the most pressing current public health concern globally. Despite the unprecedented speed at which safe and effective vaccines were developed, tested, and approved, slower than anticipated vaccine production and distribution have hampered efforts to curtail the ongoing pandemic (1–3). The delays have not only resulted in an alarming number of potentially preventable deaths but also, contributed to the emergence of more virulent strains that could reduce the efficacy of current vaccines (4). Of the eight COVID-19 vaccines currently authorized around the globe, most have a recommended schedule of two doses separated by 3 to 4 wk. While awaiting replenishment of vaccine supplies, some governments are considering or have implemented delayed administration of the second vaccine dose in order to increase the proportion of the population with at least partial protection from a single dose. Recent debates argue in favor of such delay strategies (1–3), but there are no guidelines as to how the optimal delay duration should be determined, and this has yet to be tested in clinical trials. Two major questions are the extent to which delaying the second dose would impact hospitalization rates and alleviate mitigation measures when compared with the standard 3- to 4-wk second-dose delay.

Here, we report a model-based strategy for identifying the optimal delay time between vaccine doses using intensive care unit (ICU) hospitalizations as the key metric. We show that the optimal time is influenced by both the first-dose efficacy and whether the vaccine prevents infection or alleviates symptoms. The decision to delay the second vaccine dose is therefore not trivial because single doses of infection-blocking and disease-modifying vaccines are unlikely to have an equivalent impact on curbing the pandemic (5), and a large number of infection scenarios and their effect on disease evolution must also be taken into account.

To assist in decision-making about the optimal time to delay the second COVID-19 vaccine dose, we extended the susceptible, exposed, infectious, and removed (SEIR) model to include vaccination campaigns with two doses across populations. The epidemiological dynamic is used to predict ICU bed occupancy as a stochastic process that follows the disease trajectory. This information feeds a larger model that maintains the ICU occupancy within the operational capacity by restricting population circulation when needed. This is a unique feature of our approach, as other efforts to determine optimal dosing schedules for the COVID-19 vaccines do not take into account social distancing (6–8). We describe an optimization algorithm to find the best vaccination regimen that does not increase hospitalization rates while successfully alleviating social distancing and travel restrictions, thereby accelerating a return to "normality" as soon as possible.

1. Results

The algorithm explores multiple scenarios to obtain the best profile for delaying the second vaccine dose while safeguarding

Significance

Shortages of COVID-19 vaccines hampered efforts to fight the current pandemic, leading experts to argue for delaying the second dose to provide earlier first-dose protection to twice as many people. We designed a model-based strategy for identifying the optimal second-dose delay using the hospitalization rate as the key metric. While epistemic uncertainties apply to our modeling, we found that the optimal delay was dependent on first-dose efficacy and vaccine mechanism of action. For infection-blocking vaccines, the second dose could be delayed $\geq$8 weeks if the first-dose efficacy was $\geq$50%. For symptom-alleviating vaccines, this delay duration is recommended if the first-dose efficacy was $\geq$70%. These results suggest that delaying the second vaccine dose is a feasible option.

Author contributions: P.J.S.S., C.S., C.J.S., and T.P. designed research; P.J.S.S., C.S., C.J.S., and T.P. performed research; C.S. contributed new reagents/analytic tools; P.J.S.S., C.S., L.G.N., C.J.S., and T.P. analyzed data; and P.J.S.S., C.S., C.J.S., and T.P. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

[1]To whom correspondence may be addressed. Email: tiago@icmc.usp.br.

Published August 18, 2021.

Downloaded from https://www.pnas.org by 177.185.39.185 on April 9, 2026 from IP address 177.185.39.185.

the health care system and lessening mitigation measures. The second-dose delay mechanism is not explicit but emerges from the solution of the optimization, and the solver decides when to administer the second dose. As illustrated in Fig. 1, several factors are taken into account: age, mechanism of vaccine action (symptom alleviation vs. infection prevention), vaccine availability, ICU capacity and its variability, and ICU hospitalization rates among different age groups.

To illustrate, consider a single age group. At each incremental change in delay time, there are three decision variables: 1) the reproduction number r_t, 2) the number of first doses currently administered, and 3) the number of second doses currently administered. Reducing mitigation measures (i.e., increasing r_t) increases the number of ICU admissions, whereas increasing the vaccination rate has the opposite effect. However, a certain time delay between first and second doses may be beneficial depending on the protection level conferred by the first dose. Each iteration of the solution method adjusts the control variables taking into account the dynamics, the logistic constraints, and the ICU capacity. The adjustment is performed in a manner that favors large reproduction numbers r_t. We can then compute a second-dose delay for all time steps to obtain the optimal delay.

1A. Multiple Subpopulation SEIR Model with Two Vaccine States. We extend the SEIR model to include individuals receiving vaccines.

Given a subpopulation p, the proportions of the susceptible subpopulation that have received zero, one, and two vaccine doses are $S_{p,1}, S_{p,2}$, and $S_{p,3}$, respectively. The other SEIR states, E, I, and R, were also similarly split. The parameters $a_1, a_2 \in [0, 1]$ decrease the probability that a vaccinated individual will be infected when exposed to an infected individual, considering one and two doses, respectively, while $v_{p,1}$ and $v_{p,2}$ represent the proportions of individuals who have received first and second doses. The model is presented in *Methods*, and the interactions are illustrated in Fig. 1. The main elements of our optimization model are described below.

1A.1. Groups in the population. We stratify the population into subgroups with specific demands on health care. We considered four age groups of 0 to 19, 20 to 49, 50 to 64, and 65 to 90 y in groups 1, 2, 3, and 4, respectively. The contact matrix C between age groups is presented in *Methods*. We set the initial condition as $S_0 = 68.5\%$, $E_0 = 1.0\%$, $I_0 = 0.6\%$, and $R_0 = 29.9\%$, corresponding to typical values for multiple countries at the beginning of February 2021. The algorithm was run with basal reproduction number r_0 ranging between 1.8 and 3.0 without observing a significant change in the output decisions. The subpopulation breakdown by age is 30, 48, 14, and 8% for groups 1, 2, 3, and 4, respectively.

Given a reproduction number r_0, the reproduction number of the subpopulation p is the product $b_p r_0$, where factor b_p represents subpopulation susceptibility that is greater than or equal to that of the overall population. This factor is one except for $b_2 = 1.3$, which takes into account the increased mobility and thus, contribution to disease transmission by the 20- to 49-y-old age group.

1A.2. Objectives. To prevent collapse of the health care system, the reproduction number must be decreased. This can be achieved by restricting circulation and implementing a vaccination campaign: that is, deciding for each day t and subpopulation p the values of $v_{p,1}(t)$ and $v_{p,2}(t)$ (9, 10) until the reproduction number reaches a target value $r(t)$.

Rather than performing simulations of known case studies defined by experts (1, 2) in our model, the optimization algorithm

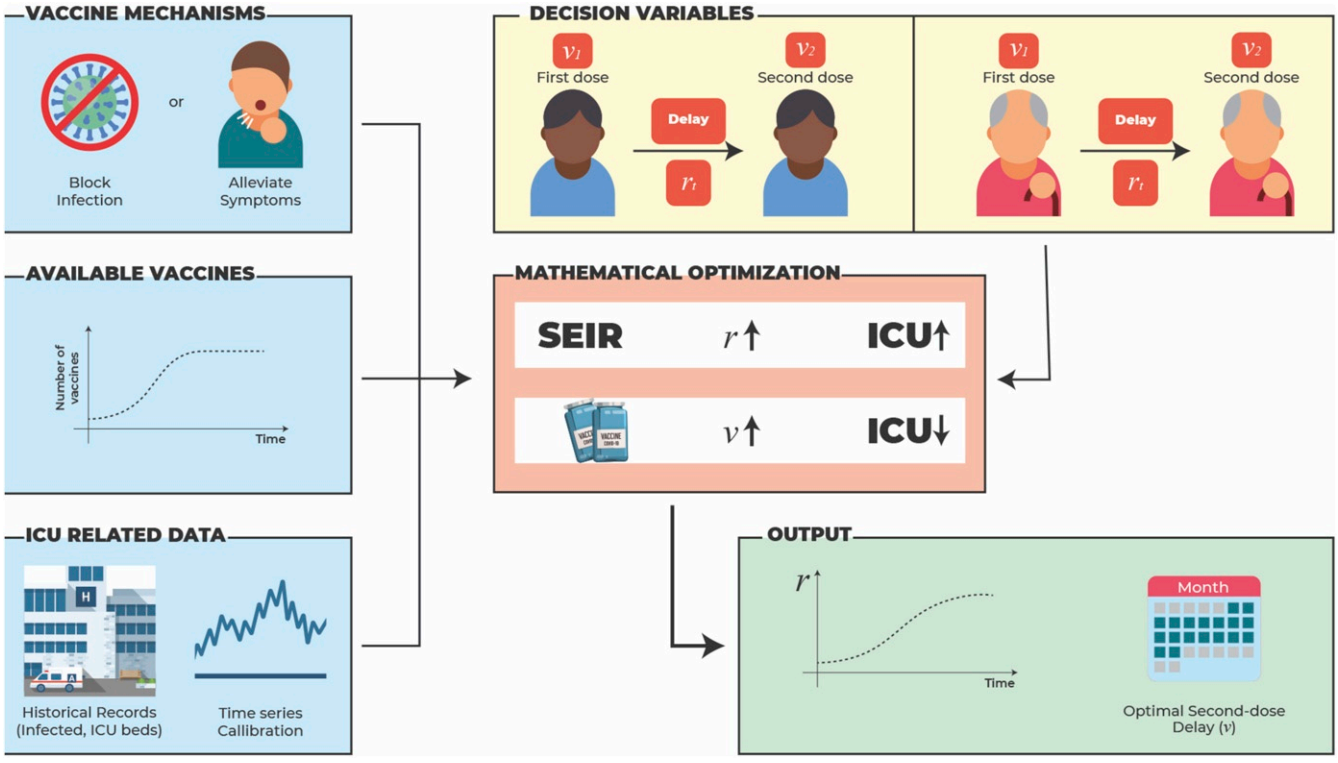

Fig. 1. A mathematical model combines data and optimization to obtain an optimal delay for administration of a vaccine second dose. The optimal second-dose delay emerges from the solution of the optimization model. The model blends ICU usage data, the vaccine mechanism of action, vaccine availability, and population demographics into an epidemiological model to predict future ICU admissions by age group. The model is solved using an optimization algorithm that considers multiple scenarios and iteratively adjusts the decision variables to find the optimal delay between the first and second vaccine doses and the target control reproduction number. The graphics used to prepare Fig. 1 are extracted from Freepik.

Downloaded from https://www.pnas.org by 177.185.39.185 on April 9, 2026 from IP address 177.185.39.185.

drives selection of the best strategy by considering implicitly all possible combinations of social distancing and vaccination profiles. Accordingly, the optimization model casts $r(t)$, $v_{p,1}(t)$, and $v_{p,2}(t)$ as decision variables to be selected by the algorithm among all scenarios that combines mitigation measure and vaccine administration to save ICU admissions. The choices range from complete lockdown $(r(t)=0)$ to free circulation $(r(t)=r_0)$ in combination with single-dose administration to any or no portion of a population $(v_{p,1}(t)\in[0,1])$, and similarly for the second dose $(v_{p,2}(t)\in[0,1])$. To determine the best strategy, the optimization process is guided by its objective function that encourages higher values of $r(t)$ and translates into a more freely circulating population. For logistic reasons, the objective function also includes a term that promotes smooth changes in the vaccination profile.

The optimization model also takes into account ICU occupancy and maintains it below maximum capacity with a 95% probability. The number of ICU patients on day t follows a stochastic process calibrated as a time series for the proportion of infected individuals who will need intensive care on day t. We estimate that the average stay in the ICU is 7 d (11). Using official records for Brazil for the last quarter of 2020, we computed the mean number of patients in the ICU for each day t between day $t-7$ and day t and divided that by the number of new cases reported in day t. After the time series is calibrated for the ratio, knowing the number of infected individuals in the SEIR dynamics at day t gives a stochastic trajectory of ICU bed occupancy from which we derive a probabilistic constraint. The considered ICU bed capacity is 17.5 per 100,000 inhabitants, and the mean ICU bed demand is 1.2% (SD 0.64%) of the infected population, distributing the values among the age groups.

1A.3. Vaccine mechanisms. We consider two types of vaccine (5): infection blocking and symptom alleviating. An infection-blocking vaccine can prevent infection by decreasing the likelihood of becoming infected from an encounter with an infectious individual by one minus the dose efficacy. This is modeled by multiplying the respective reproduction number by this factor, represented by the parameters a_1 and a_2 in the model. In the case of a vaccine that alleviates symptoms, we keep $a_1=a_2=1$ and assume that the effect of the vaccine is only to decrease the likelihood of an infected individual requiring ICU admission. This is modeled by multiplying the estimate of ICU beds needed for vaccinated individuals by one minus the dose efficacy.

1A.4. Profile of available doses. Available dose numbers are given as a proportion of the population size, with an initial number of 0.1% ramping up linearly to 1.0% over 150 d and remaining constant thereafter. Small modifications to this profile do not affect the main results.

1B. Standard Vaccination by Age Groups as Baseline. For both types of vaccine mechanisms over a period of 364 d, we first performed a simulation in which the standard two-dose vaccination strategy is administered to the most at-risk age group (65 to 90 y) first. As a "baseline" strategy, the second dose is automatically reserved for administration 4 wk later, and subpopulations are parsed in order of decreasing age. To evaluate the benefits of postponing the second dose, we used the number of ICU admissions as a measure of benefit gained by optimization of the vaccination strategy compared with the baseline strategy. A complementary metric is to assess the closeness of the reproduction number r_t to the maximum value $r_0 = 2.5$. Since a value of 2.5 represents no circulation restriction, closeness to r_0 reflects proximity to an "open society" without enforced mitigation measures.

1C. Delaying the Second Dose. To address whether delaying the second vaccine dose is desirable, the optimization model determines the best combination of target reproduction number $r(t)$ and vaccine distribution among the age groups. In our model, the second dose can be administered at any time up to 12 wk after the first dose.

1C.1. Impact of first-dose efficacy and the vaccine mechanism. We fixed the second-dose efficacy at $a_2=82.4\%$, as observed for the AstraZeneca Oxford vaccine (AZD1222) (12), and let the first-dose efficacy a_1 vary from 0 to 80% in steps of 5%. The goals are 1) to determine if an optimal delay exists (other than the standard 4 wk) and 2) to determine if a threshold exists that would trigger a switch in the best delay time from the recommended (i.e., 4 wk) to as late as possible (12 wk).

The results for infection-blocking and symptom-alleviating vaccines are shown in Fig. 2, where the dotted lines mark the delay at $r_0=2.5$ and the whiskers correspond to all possible values when r_0 is varied from 1.8 to 3 in 0.2 steps. For r_0 across this range, the threshold efficacy to obtain herd immunity (defined as the point at which the effective reproduction number is below one) ranges from 44.4 to 66.6%. For example, for $r_0=1.8$, the model predicts that a fully open society can be reached by week 5, well before the maximum allowed delay for the second dose

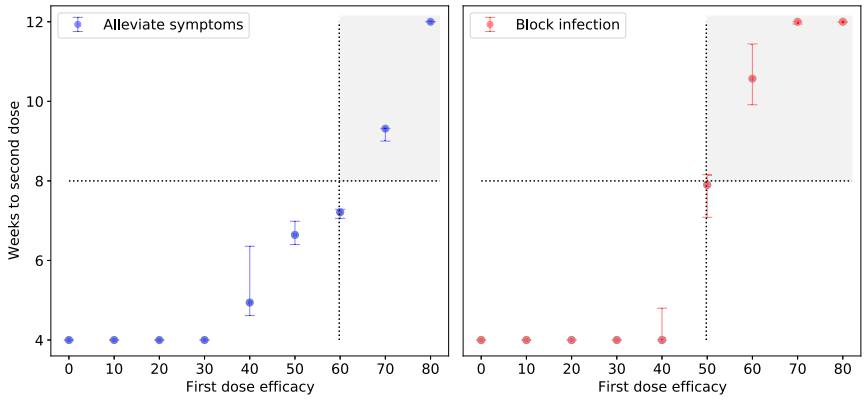

Fig. 2. Optimization of the delay between the first and second doses of a COVID-19 vaccine based on first-dose efficacy and vaccine mechanism. The shaded areas represent the first-dose efficacy that results in doubling the time to second dose from the baseline (4 wk). *Left* shows the second-dose delay when the vaccine alleviates symptoms; in this case, the best strategy delays the second dose for ≥8 wk when the first-dose efficacy is ≥70%. *Right* shows the second-dose delay when the vaccine blocks infection; here, the best strategy delays the second dose for ≥8 wk when the first-dose efficacy is ≥50%. For both vaccine types, the second-dose efficacy reaches 82.4%. The filled circles show the time to the second dose for $r_0=2.5$, and the bars represent the variability across simulations when r_0 is varied from 1.8 to 3 in 0.2 steps.

Downloaded from https://www.pnas.org by 177.185.39.185 on April 9, 2026 from IP address 177.185.39.185.

(12 wk). In contrast, for $r_0 = 3.0$, the model predicts that mitigation measures will still be required beyond 12 wk. The decision to postpone the second dose is robust across the considered reproduction numbers. These results show that the vaccine mechanism plays a major role in the decision. Thus, when the vaccine blocks infections, the algorithm suggests that the second dose can be delayed by $\geq$8 wk if the first-dose efficacy is $\geq 50\%$ and by the maximum delay of 12 wk if the first-dose efficacy is $\geq 70\%$, whereas a delay ≥ 8 wk is advisable for symptom-alleviating vaccines only when the first-dose efficacy is $\geq 70\%$.

1C.2. Effect of augmented transmission in a subpopulation. Individuals within specific age groups may play a more active role than other age groups in spreading the disease due to a higher degree of social activity or less stringent attention to sanitary measures. For this analysis, we considered people in the 20- to 49-y-old age group, most of whom would be working and socially active. We fix r_0 at 2.5 and compute the second-dose timing as a function of the first-dose efficacy for b_2 ranging from 1 to 1.6. We observed only a small variability in the optimal time to second dose; in fact, all curves fell within the shaded area of Fig. 2. Thus, postponing the second dose is robust with respect to changes in both r_0 and the reproduction number of specific subpopulations.

1C.3. Comparison between the optimally delayed second dose and baseline strategy. We let $r_0 = 2.5$ for all age groups except the 20- to 49-y-old group, where r_0 was set 30% higher. We fixed the first- and second-dose efficacies at 76 and 82.4%, respectively, as estimated for the AZD1222 vaccine (12).

To quantify the effect of delaying the second dose on health, we computed the difference in ICU demand when comparing the optimization model and baseline strategies. We varied the efficacy of the first dose from 0 to 80% and determined the impact on COVID-19–related ICU admissions over a specific period, which we set here as 200 d. The algorithm finds an optimal second-dose delay in which social distancing values are no worse than those with the baseline strategy. To achieve this, the model includes the mean baseline r_t as a lower bound and runs the optimization problem with a metric that decreases the overall ICU usage.

The results for the first 200 d are reported in Fig. 3. The data show that a second-dose delay beyond the standard 4 wk will result in a progressive reduction in ICU admissions as the first-dose efficacy increases, with a threshold of 50%. As a comparison, this model would predict that for an infection-blocking vaccine such as AZD1222, which has an actual first-dose efficacy of about 76%, an optimized second-dose delay would save 45 ICU admissions per 100,000 individuals compared with the baseline strategy.

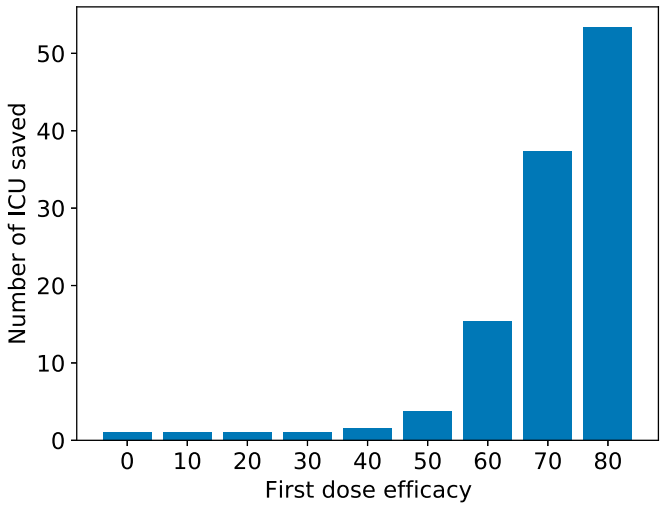

Fig. 3. Reduction in ICU occupancy using the optimized second-dose delay strategy compared with the standard delay strategy. The graph shows the predicted reduction in ICU admissions when the first-dose efficacy is varied from 0 to 80% in 5% increments and is normalized to the baseline strategy of 4 wk between doses.

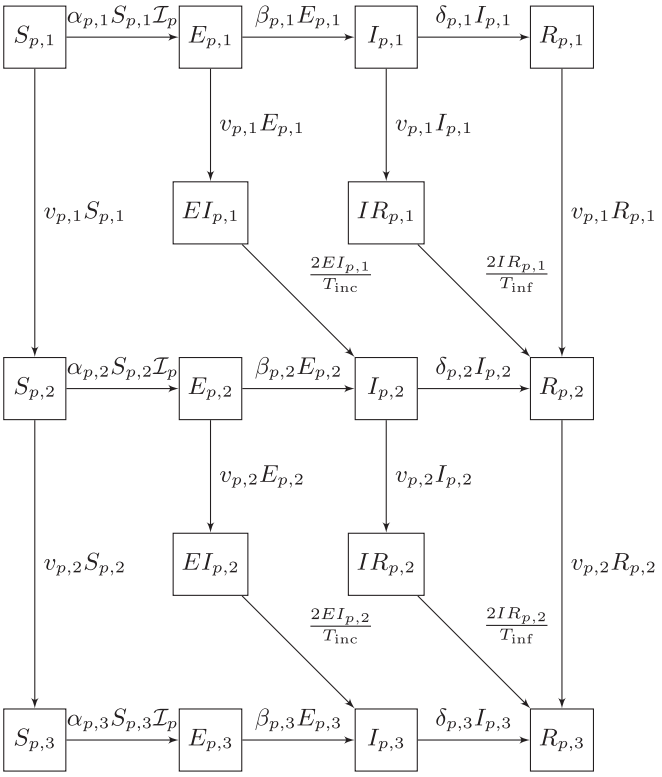

Fig. 4. SEIR model for age group *p* with a two-dose vaccine that blocks infection.

Assuming a first-dose efficacy of $\geq$50%, the optimization model indicates that each 6% increase in first-dose efficacy reduces the number of ICU admissions by 10 per 100,000 individuals for infection-blocking vaccines and 8 per 100,000 individuals for symptom-alleviating vaccines.The optimized second-dose delay strategy is also supported by complementary metrics that examine the extent to which the vaccination campaign leads progressively to a fully open society. Whereas the baseline strategy results in three periods of 2 wk each of strict lockdown, an optimized second-dose delay of 12 wk results in a single initial lockdown of 2 wk followed by a gradual relaxation of social distancing measures. Thus, the optimization model predicts that a second-dose delay would be beneficial not only in reducing the burden on the health care system but also, in accelerating the return of society to a normal lifestyle.

2. Discussion

We developed a model to examine the impact of optimizing the delay between doses of a two-dose COVID-19 vaccine on ICU admissions and societal restrictions compared with a standard delay of 4 wk. Our study demonstrates that the mechanism of vaccine action has a profound effect on the delay. Thus, if the vaccine blocks infection with a first-dose efficacy of $\geq$50%, the second dose can be delayed by at least 8 wk after the first dose (half of the maximal delay allowed). In contrast, the same delay would be advisable for a symptom-alleviating vaccine only if the first-dose efficacy is at least 70%.

Downloaded from https://www.pnas.org by 177.185.39.185 on April 9, 2026 from IP address 177.185.39.185.

Strategies to extend the reach of available vaccine supplies, such as halving the dose or increasing the delay interval between doses as considered here, could have a significant beneficial effect on disease transmission and mortality until vaccine supplies increase (1, 2). Critics of these strategies fear that induced immunity may wane before the second dose can be administered or result in weak immunity that would favor the emergence of variant strains with higher virulence (13). However, even partial immunity may reduce disease and/or infection and concomitantly lessen the opportunity for viral mutation and propagation (14–17).

Our findings suggest that several societal benefits can be obtained by delaying the second dose of a vaccine, especially if the first-dose efficacy is high (e.g., 76% for AZD1222). First, an optimized delay could lead to a massive reduction in the number of ICU admissions. Given that ICU mortality rates for COVID-19 have reached about 60% in some countries (18), our findings suggest that an optimized second-dose delay for infection-blocking and symptom-alleviating vaccines could save 90,000 and 40,000 lives, respectively, over 200 d in a country with 300 million inhabitants. The second benefit obtained by an optimized second-dose delay is the extent to which it reduces mitigation measures. Our model predicts a reduction in full lockdowns from three 2-wk lockdowns when using the standard second-dose delay of 4 wk to a single 2-wk lockdown followed by a gradual social distancing relaxation with the optimized delay. Thus, a second-dose delay can save lives and accelerate the return to free circulation.

3. Methods

3A. SEIR Model with Multiple Doses of Vaccine. First, we discuss the SEIR dynamic when a two-dose vaccination campaign is put in place. For a fixed population p and a vaccine mechanism that blocks transmission, we consider the diagram represented by Fig. 4.

The mean incubation and infectious times are T_{inc} and T_{inf}, respectively. The variable $S_{p,1}$ represents the proportion of the subpopulation p that is in the first stage of the vaccination process: that is, susceptible individuals in age group p who did not receive any dose of the vaccine yet or who have been recently inoculated and the first-dose immunity is still building up. Likewise, $S_{p,2}$ is the proportion of susceptible individuals who are already protected by the effect of the first dose while still waiting for the second dose to be inoculated or become effective. Finally, $S_{p,3}$ represents the fraction of susceptible already protected by the full effect of the two doses. The meaning of the other compartments, $E_{p,i}$, $I_{p,i}$, $R_{p,i} = 1, \ldots, 3$, is similar. The intermediate states $EI_{p,i}$ and $IR_{p,i}$ represent individuals in $E_{p,i-1}$ and $R_{p,i-1}$ for whom the ith vaccine dose has just made effect and who are transitioning to the new states $I_{p,i}$ and $R_{p,i}$, respectively. Note that they spend only half of the time in this state, as they have already spent some time in the original state before moving to the next stage.

The variable $v_{p,i}$, $i = 1, 2$, models the number of people from the level $i-1$ for which the ith dose of the vaccine made effect at a given moment in time, afterwards their state changes and they move to the next level i. Other transition relations can be found in Table 1, where a_1, a_2 stands for the level achieved in attenuating the infection after receiving one and two doses of the vaccine. The factor b_p models the effect that behavioral patterns of the subpopulation p have on the chances of getting infected when encountering infectious individuals (for instance, younger and active individuals being more prone to agglomerate have an increased risk of infection). The variable $r(t)$ denotes the target reproduction number that is needed at time t to control the spread of the disease with nonpharmaceutical measures, such as social distancing or mask wearing. Also, we note that the diagram in Fig. 4 has no transitions from $EI_{p,1}$ and $IR_{p,1}$ to the lower levels because we are assuming that the meantime in these states, $T_{\text{inc}}/2$ and $T_{\text{inf}}/2$, is significantly lower than the time required for a vaccine dose to make an effect.

Table 1. Transition ratios for levels in the diagram in Fig. 4

	$\alpha_{p,i}$	$\beta_{p,i}$	$\delta_{p,i}$
$i=1$	$\frac{b_p r_t (1-v_{p,1})}{T_{\text{inf}}}$	$\frac{1-v_{p,1}}{T_{\text{inc}}}$	$\frac{1-v_{p,1}}{T_{\text{inf}}}$
$i=2$	$\frac{a_1 b_p r_t (1-v_{p,2})}{T_{\text{inf}}}$	$\frac{1-v_{p,2}}{T_{\text{inc}}}$	$\frac{1-v_{p,2}}{T_{\text{inf}}}$
$i=3$	$\frac{a_2 b_p r_t}{T_{\text{inf}}}$	$\frac{1}{T_{\text{inc}}}$	$\frac{1}{T_{\text{inf}}}$

Table 2. ICU demand by age for the state of São Paulo, Brazil

Population group, p	Age group, y	Demography, %	Actual ICU usage, %	Correcting factor, γ_p
1	0–19	29	2	0.06
2	20–49	48	28	0.58
3	50–64	14	30	2.06
4	65–90	8	40	5.16

Finally, the variable $\mathcal{I}_p$ binds all the subpopulations together by a *contact matrix* that describes how the different age groups interact. This is a square matrix C of dimension #subpopulations × #subpopulations in which the entry $C_{p,p'}$ represents the proportion of contacts the subpopulation p makes with individuals of the subpopulation p'. In particular, the sum by rows equals 1.0. Using this matrix, we define

$$I_p := \frac{I_{p,1} + I_{p,2} + I_{p,3} + IR_{p,1} + IR_{p,2}}{\text{proportion of the overall population that belongs to } p}.$$

Additionally,

$$\mathcal{I}_p = \sum_{p'=1}^{\#\text{populations}} C_{p,p'} I_{p'},$$

for which the contact matrix is

$$C = \begin{pmatrix} 0.57 & 0.27 & 0.10 & 0.06 \\ 0.20 & 0.59 & 0.15 & 0.06 \\ 0.15 & 0.46 & 0.27 & 0.12 \\ 0.18 & 0.24 & 0.18 & 0.39 \end{pmatrix}.$$

3B. ICU Occupation. Since the percentage of the infected population that needs intensive care at time t is represented by a stochastic process, a probabilistic constraint keeps the ICU bed usage below the maximum capacity, $v_t^{\max}$. The use of a time series for the stochastic process makes it possible to reformulate the probabilistic constraint into an equivalent deterministic inequality. The detailed procedure is explained in ref. 19; here, we mention a few key points only.

Suppose infected individuals spend on average $\nu = 7$ days in intensive care (20). At time t, if the bed usage for a given age group p is v_p^t, then the fraction of infected individuals who need ICU beds at time t can be estimated by the ratio $v_p^t / \mathcal{J}_p^t$, where $\mathcal{J}_t^p$ corresponds to individuals in the subpopulation p moving to the compartment R at time t. In our setting, this is equal to $1/T_{\text{inf}} \left(\sum_{k=t-\nu}^{t} \sum_d (I_{pd}^k + 2IR_{pd}^k) \right)$. The total ICU demand is distributed proportionally among the age groups using a correcting factor γ_p that represents the proportion of the age group that uses ICU beds, when compared with its demography. Table 2 reports the corresponding data for 2010 in São Paulo state, which has 44 million inhabitants.

To obtain the time series parameters, we computed the history of ratios from the last quarter of 2020 until the end of January 2021 using official records of São Paulo state. Autoregressive models with lag up to two and no trend appeared as the best ones in terms of Bayesian information criterion. Back testing over January gave mean absolute percentage error values of 1.88, 2.02, and 9.84% for the lags two, one, and zero, respectively, with an approximate normal-looking shape for the residual histograms. When performing out of sample simulations until May 2021, we noticed that both trajectories with positive lags stabilized at the same value as those with zero lag. Since the mean absolute percentage error accuracy of the latter model was acceptable and the optimization problem covers a long time horizon, we chose the simplest time series, with lag zero, for the constraint. After scaling back, this procedure gave the following simple model, $\text{icu}^t(\omega) \sim \mathcal{N}(c_0, \sigma_\omega^2)$ for $c_0 = 0.01$ and $\sigma_\omega = 0.002$. After these parameters are known, the pro rata per subpopulation is done as explained above to make explicit the probabilistic constraints

$$\mathbb{P}\left[\text{icu}^t(\omega) \sum_{p=1}^{4} \gamma_p \mathcal{J}_p^t \le v_{\max}^t \right] \ge \delta,$$

which ensure the hospital capacity will not be exceeded, with δ probability. In the experiments, we used $\delta = 95\%$. The deterministic equivalent of the chance constraint is an affine inequality of the form $c_0 \sum_p \gamma_p \mathcal{J}_p^t \le v_{\max}^t + \Delta$,

APPLIED MATHEMATICS

Downloaded from https://www.pnas.org by 177.185.39.185 on April 9, 2026 from IP address 177.185.39.185.

where the term Δ involves the inverse cumulative function of the standard Gaussian distribution. Ref. 19 has more details.

3C. Optimizing on Multiple Populations. Given the models for ICU demand and vaccine supplies, the optimization algorithm finds the best strategy for vaccination and social distancing that simultaneously avoids the collapse of the health care infrastructure and controls the pandemic in the shortest amount of time. To achieve these goals we initially use an objective function f that encourages large values of the reproduction number r_t. The optimization is done over the decision variables r_t, $v_{p,1}$, and $v_{p,2}$ that control the trajectories of the state variables, defined by the different SEIR compartments. The system of differential equations of the modified SEIR dynamic illustrated by the diagram in Fig. 1 is discretized using central finite differences over a horizon of K days ($K = 364$ in the simulations). Letting c denote the vector whose components are the control variables r_t^i, $v_{p,1}^i$, and $v_{p,2}^i$, for $i = 1, \ldots, K$ and $p = 1, \ldots,$ #subpopulations, and letting s denote the state vector, with all the SEIR compartments for all subpopulations, the optimization problem has the abstract form

$$\begin{aligned} \max_{c,s} \quad & f(c) \\ \text{s.t.} \quad & s \in \text{Discretization of SEIR}(c) \\ & c \in \text{logistic bounds} \\ & s \in \text{ICU usage}, \end{aligned}$$

where the latter set represents the reformulation of the probabilistic constraint defined in Section 3B. Logistic relations on the control variables refer to the number of vaccines available each day taking into account both vaccine availability and the inoculation capacity. Specific time windows for giving the second dose, say $[t_{\min}, t_{\max}]$, are enforced by requiring the total number of second doses applied up to day t to be smaller than or equal to the total number of first doses administered up to the day $t - t_{\min}$. This constraint ensures that no second doses are given before possible. On the other hand, we also limit the total number of second doses given until t to be at least the total number of first doses administered up to $t - t_{\max}$, ensuring that the second doses are applied within the desired time window. In our implementation, $t_{\min} = 28$ and $t_{\max} = 84$.

The objective function to minimize is the sum of differences $\sum_{i=1}^{K} r_0 - r_t^i$, where $r_0 = 2.5$ is the basal reproduction number of the virus used in the simulation (21, 22). The rationale is to allow the population to circulate freely. To discourage bang-bang controls, we add to the objective a total quadratic variation term for the number of doses applied at each population each day. This term results in vaccination profiles that are easier to visualize and understand. After this first optimization step, we perform a second optimization that tries to minimize ICU usage without deteriorating the computed r_t^* profile. To achieve this, we add the constraint

$$\sum_{i=1}^{K} r_t^i \geq \sum_{i=1}^{K} (r_t^*)^i,$$

and we switch the objective to minimize the estimated ICU usage.

These problems are large-scale nonconvex quadratic optimization problems. They are formulated using the JuMP (23) modeling language and solved using the nonlinear optimization solver Ipopt (24). The typical solution time is around 5 min on a desktop computer with a Ryzen 1700X processor.

Data Availability. Input files or sets of input parameters as well as the source code have been deposited in GitHub (https://github.com/pjssilva/Robot-vaccine).

ACKNOWLEDGMENTS. This work was supported by Center for Research in Mathematics Applied to Industry, Fundação de Amparo à Pesquisa do Estado de São Paulo Grants 2013/07375-0 and 2018/24293-0; the Royal Society London; Brazilian National Council for Scientific and Technological Development Grants 301778/2017-5, 302836/2018-7, 304301/2019-1, 306090/2019-0, 302912/2019-3, and 403679/2020-6; Serrapilheira Institute Grant Serra-1709-16124; and Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro Grant E-26/202.828/2017.

1. A. D. Paltiel, A. Zheng, J. L. Schwartz, Speed versus efficacy: Quantifying potential tradeoffs in COVID-19 vaccine deployment. *Ann. Intern. Med.* **174**, 568–570 (2021).
2. R. V. Barnabas, A. Wald, A public health COVID-19 vaccination strategy to maximize the health gains for every single vaccine dose. *Ann. Intern. Med.* **174**, 552–553 (2021).
3. The Lancet Infectious Diseases, An exceptional vaccination policy in exceptional circumstances. *Lancet Infect. Dis.* **21**, 149 (2021).
4. A. Fontanet *et al*., SARS-CoV-2 variants and ending the COVID-19 pandemic. *Lancet* **397**, 952–954 (2021).
5. M. Lipsitch, N. E. Dean, Understanding COVID-19 vaccine efficacy. *Science* **370**, 763–765 (2020).
6. C. R. MacIntyre, V. Costantino, M. Trent, Modelling of COVID-19 vaccination strategies and herd immunity, in scenarios of limited and full vaccine supply in NSW, Australia.*Vaccine*, 10.1016/j.vaccine.2021.04.042 (2021).
7. L. Matrajt, J. Eaton, T. Leung, E. R. Brown, Vaccine optimization for COVID-19: Who to vaccinate first? *Sci. Adv.* **7**, eabf1374(2020).
8. L. Matrajt *et al*., Optimizing vaccine allocation for COVID-19 vaccines: Critical role of single-dose vaccination. medRxiv [Preprint] (2021). https://doi.org/10.1101/2020.12.31.20249099 (Accessed 11 February 2021).
9. S. L. Chang, N. Harding, C. Zachreson, O. M. Cliff, M. Prokopenko, Modelling transmission and control of the COVID-19 pandemic in Australia. *Nat. Commun.* **11**, 5710 (2020).
10. J. Dehning *et al*., Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science* **369**, eabb9789 (2020).
11. S. Aziz *et al*., Managing ICU surge during the COVID-19 crisis: Rapid guidelines. *Intensive Care Med.* **46**, 1303–1325 (2020).
12. M. Voysey *et al*., Single dose administration, and the influence of the timing of the booster dose on immunogenicity and efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine: A pooled analysis of four randomised trials. *Lancet* **397**, 881–891 (2021).
13. P. Bieniasz, The case against delaying SARS-CoV-2 mRNA vaccine boosting doses. *Clin. Infect. Dis.*, ciab070 (2021).
14. S. Cobey, D. B. Larremore, Y. H. Grad, M. Lipsitch, Concerns about SARS-CoV-2 evolution should not hold back efforts to expand vaccination. *Nat. Rev. Immunol.* **21**, 330–335 (2021).
15. D. A. Kennedy, A. F. Read, Why the evolution of vaccine resistance is less of a concern than the evolution of drug resistance. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 12878–12886 (2018).
16. D. A. Kennedy, A. F. Read, Why does drug resistance readily evolve but vaccine resistance does not? *Proc. Biol. Sci.* **284**, 20162562 (2017).
17. C. M. Saad-Roy *et al*., Epidemiological and evolutionary considerations of SARS-CoV-2 vaccine dosing regimes. *Science* **372**, 363–370 (2021).
18. O. T. Ranzani *et al*., Characterisation of the first 250,000 hospital admissions for COVID-19 in Brazil: A retrospective analysis of nationwide data. *Lancet Respir. Med.* **9**, 407–418 (2021).
19. L. G. Nonato, P. Peixoto, T. Pereira, C. Sagastizábal, P. J. S. Silva, Data from "Robot Dance: A mathematical optimization platform for intervention against COVID-19 in a complex network." Optimization Online. http://www.optimization-online.org/DB_HTML/2020/10/8054.html. Accessed 1 June 2020.
20. K. V. M. d. S. Noronha *et al*., Pandemia por COVID-19 no Brasil: Análise da demanda e da oferta de leitos hospitalares e equipamentos de ventilação assistida segundo diferentes cenários. *Cad. Saude Publica* **36**, e00115320 (2020).
21. D. S. Candido *et al*., Evolution and epidemic spread of SARS-CoV-2 in Brazil. *Science* **369**, 1255–1260 (2020).
22. L. Ferretti *et al*., Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* **368**, eabb6936 (2020).
23. I. Dunning, J. Huchette, M. Lubin, JuMP: A modeling language for mathematical optimization. *SIAM Rev.* **59**, 295–320 (2017).
24. A. Wächter, L. T. Biegler, On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.* **106**, 25–57 (2006).

Downloaded from https://www.pnas.org by 177.185.39.185 on April 9, 2026 from IP address 177.185.39.185.